

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Patent Application of:

Yasuyuki MIURA et al.

Application No.: Unassigned

Group Art Unit: Unassigned

Filed: August 19, 2003

Examiner: Unassigned

For: REAL-TIME CONTENTS EDITING METHOD, SYSTEM, AND PROGRAM

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN  
APPLICATION IN ACCORDANCE  
WITH THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Commissioner for Patents  
PO Box 1450  
Alexandria, VA 22313-1450

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s) herewith a certified copy of the following foreign application:

Japanese Patent Application No(s). 2003-115165

Filed: April 21, 2003

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing date(s) as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. § 119.

Respectfully submitted,

STAAS & HALSEY LLP

Date: August 19, 2003

By:   
David M. Pitcher  
Registration No. 25,908

1201 New York Ave, N.W., Suite 700  
Washington, D.C. 20005  
Telephone: (202) 434-1500  
Facsimile: (202) 434-1501

日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office

出願年月日

Date of Application: 2003年 4月21日

出願番号

Application Number: 特願2003-115165

[ST.10/C]:

[JP2003-115165]

出願人

Applicant(s): 独立行政法人通信総合研究所

2003年 5月13日

特許庁長官  
Commissioner,  
Japan Patent Office

太田信一郎



出証番号 出証特2003-3035490

【書類名】 特許願  
【整理番号】 CRL-03-39  
【提出日】 平成15年 4月21日  
【あて先】 特許庁長官殿  
【国際特許分類】 G06F 15/00  
【発明者】  
【住所又は居所】 東京都小金井市貫井北町4-2-1 独立行政法人通信  
総合研究所内  
【氏名】 三浦 康之  
【発明者】  
【住所又は居所】 東京都小金井市貫井北町4-2-1 独立行政法人通信  
総合研究所内  
【氏名】 勝本 道哲  
【特許出願人】  
【識別番号】 301022471  
【氏名又は名称】 独立行政法人通信総合研究所  
【代理人】  
【識別番号】 100108660  
【弁理士】  
【氏名又は名称】 大川 譲  
【電話番号】 03-3807-1151  
【選任した代理人】  
【識別番号】 100097836  
【弁理士】  
【氏名又は名称】 福井 國敞  
【提出物件の目録】  
【物件名】 明細書 1  
【物件名】 図面 1  
【物件名】 要約書 1

特2003-115165

【包括委任状番号】 0201220

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 実時間コンテンツ編集方法及びシステム並びにプログラム

【特許請求の範囲】

【請求項1】 インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集方法において、

入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の1つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えて、

配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンVOPの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する、

ことから成る実時間コンテンツ編集方法。

【請求項2】 前記ビデオオブジェクトプレーンVOPの種類および使用頻度の変更は、入力画像の符号化処理を基本処理と付加的処理に分類し、分類した各種処理の符号化実行時間を計測し、その計測結果に基づいて行う請求項1に記載の実時間コンテンツ編集方法。

【請求項3】 インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集システムにおいて、

入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の1つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールから構成されて、

配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンVOPの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する、

ことから成る実時間コンテンツ編集システム。

【請求項4】 前記マシンの性能は、プロセスの実行時間の監視を行うことによって行い、監視の結果に基づいて複数の符号化アルゴリズムから適切なものを選択する請求項3に記載の実時間コンテンツ編集システム。

【請求項5】 前記プロセスの実行時間の監視は、システム内でビデオオブジェクトプレーンVOPの符号化時間を測定し、あらかじめ定められた平均フレームレートにもとづいてフレーム間圧縮を行うか否かを選択する請求項4に記載の実時間コンテンツ編集システム。

【請求項6】 前記符号化規格は、MPEG-4規格である請求項3に記載の実時間コンテンツ編集システム。

【請求項7】 前記編集モジュールは、配信サーバに対する画像及び／又は音声のマルチキャスト要求を行い、クライアントへ送信するシーン記述言語の生成およびマルチキャストを担当する請求項3に記載の実時間コンテンツ編集システム。

【請求項8】 前記選択された符号化アルゴリズムによる符号化処理を段階的に行い、一定時間経過後に必要最小限の符号化を完了し、その後に解像度や圧縮率を高めた付加的な符号化処理を行い、制限時間がすぎた時点で付加的な符号化処理が完了しなかった場合、その時点で処理を打ち切り、直前の段階における符号化処理の結果を配信する請求項3に記載の実時間コンテンツ編集システム。

【請求項9】 インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集プログラムにおいて、

入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の1つを選択して符号化可能な符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えて、

配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンVOPの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する。

各手順を実行することから成る実時間コンテンツ編集プログラム。

**【発明の詳細な説明】****【0001】****【発明の属する技術分野】**

本発明は、インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集方法及びシステム並びにプログラムに関する。

**【0002】****【従来の技術】**

近年の広帯域なネットワークの普及により、大容量の動画像コンテンツの配信が可能となっている。2001年以降、ADSL(Asymmetric Digital Subscriber Line)の加入者が大幅に増加したことにより、ブロードバンド環境が整いつつある。現在（2002年3月現在）、日本では、7.3%の世帯がADSLまたはCATVに加入し、1～10Mbps程度のデータ受信および数百キロbps程度の送信が可能となっている。また、今後FTTH(Fiber To The Home)の普及が見込まれており、これにより最大で100Mbps程度の通信が可能となり、ユーザ同士で大容量の動画像を送受信することが可能になると予想される。

**【0003】**

それにともない、広帯域ネットワークを通じて画像・音声を配信するサービスやツールが数多く提案されている。広帯域を使用した高画質映像配信システムとして、Ruff Systems（非特許文献1参照）、DVTS（非特許文献2参照）などのシステムが提案されている。Ruff Systemsは、非圧縮のD1/HDTV映像やビデオカメラから取得したDVストリームを、TCP/IP上で配信するシステムである。非圧縮画像やDV画像を使用しているため大きな帯域を必要とするが、高画質の配信が可能という特徴を有する。狭帯域に対応したMPEG動画やH.323規格を利用しシステムやツールは多数存在する（非特許文献3～6参照）。これらはいずれもPC上からカメラ画像をインターネット経由で受信することが可能であるが、単一の動画を対象としたアプリケーションであり、多数の動画を扱うには不向きである。

**【0004】****【非特許文献1】**

<http://www.cnd.tel.co.jp/product/syo64.html>

【非特許文献2】

<http://www.sfc.wide.ad.jp/DVTS>

【非特許文献3】

ohphone: [http://www.openh323.org/docs/ohphone\\_man.html](http://www.openh323.org/docs/ohphone_man.html)

【非特許文献4】

FFmpeg: <http://ffmpeg.sourceforge.net/>

【非特許文献5】

Coriander: <http://www.tele.ucl.ac.be/PEOPLE/DOUXCHAMPS/ieee1394/coriander/>

【非特許文献6】

三浦康之, 勝本道哲 「実時間環境におけるビジュアル符号化のための実証実験」,  
情報処理学会研究報告, 2002-DPS-100, pp.25-30, 2002

【0005】

【発明が解決しようとする課題】

ビデオカメラから取得した映像をその時点で配信するライブ配信を可能とするためには、符号化処理を実時間処理で行うことが求められる。実時間処理とは、外部からの連続的な入力に対し、定められた時間内に出力を返すことが要求される処理である。一秒間に  $f$  フレームの動画を配信する実時間処理システムでは、1 フレームに対する編集処理およびビットストリームのパケット化を平均  $1/f$  秒以内に完了し、その場で配信しなければならない。

【0006】

上記のような目標を達成するためには、以下のアプローチが有効である。

- ・ 符号化処理を高速化し、制限時間以内に符号化処理を完了する。
- ・ 性能の異なるマシンの上で正しく動作するために、プロセスの実行時間の監視を行い、監視の結果に基づいて複数の符号化アルゴリズムから適切なものを選択する。
- ・ 符号化処理を段階的に行う。一定時間経過後に必要最小限の符号化を完了し、その後に解像度や圧縮率を高めた付加的な符号化処理を行う。制限時間がすぎた

時点で付加的な符号化処理が完了しなかった場合、その時点で処理を打ち切り、直前の段階における符号化処理の結果を配信する。

## 【0007】

上記の考え方に基づいて、本発明の実時間コンテンツ編集システムでは、システム内でVOP（ビデオオブジェクトプレーン：Video Object Plane、フレームに相当）の符号化時間を測定し、あらかじめ定められた平均フレームレートにもとづいてフレーム間圧縮を行うか否かを選択する。

## 【0008】

本発明者は個々のストリームを対象とした配信システムに対して、複数の画像や音声を入力とし、リアルタイムで編集・コード化して配信する作業を行う実時間コンテンツ編集システムを提案する。本システムは、MPEG-4のビデオオブジェクト（VO）および編集用言語の符号化および復号を分散環境において処理するもので、従来の配信システムに対して以下のような特長を有する。

- ① 離れた複数地点からのライブ映像を用いたコンテンツ編集作業が可能
- ② 各地点からの映像配信者や映像受信者は、システムの稼動中に自由に参加・離脱が可能

個々のユーザのニーズに沿ったコンテンツを提供するためには、ユーザ側に多数のビデオオブジェクト（VO）を送りコンテンツを構築するため、ユーザ側に多数のビデオオブジェクト（VO）が集中する。そのため個々のビデオオブジェクト（VO）に多くの帯域を割くことが困難となる。また、送信者がADSLのような非対称な通信サービスを使用する場合、上り側の帯域幅は下り側のそれに比べて小さなものにとどまるため、やはり ビデオオブジェクト（VO）に多くの帯域を割くことが困難となる。Ruff Systemsのように、D1非圧縮データやDV圧縮データ等の大きなデータをそのまま伝送する方式を採用した場合、必要帯域が大きいために多数のビデオオブジェクト（VO）を同時に送受信するシステムでは莫大な量の帯域が必要となる。そのため、高い圧縮率を持ち、さまざまな帯域に対応したMPEG-4ビデオオブジェクト（MPEG-4 VO）等への符号化が必要になるが、MPEG-4 VO符号化にはフレーム間圧縮に多大なCPUパワーを必要とするため、時間的制約の厳しい実時間環境下でスペックの低いマシンを用いてライブ映像の符号化を実行す

る場合、複雑な符号化は困難となる。個人ユーザが自宅から、あるいは持ち運び可能な程度の機材を用いて出張先等から、ライブ映像を送信する等の場合、家庭用PC程度の機材しか使用できないことが多い。他方、比較的高性能なサーバマシンを使用した場合、スペックの低いマシンを使用する場合に比べ、複雑な符号化が可能になる。このように、使用するサーバの性能により可能な処理の範囲が変わってくる。

#### 【0009】

そこで、本発明は、係る問題点を解決して、CPUパワーに応じた符号化処理を行うことで、リアルタイムなMPEG-4 V0への符号化を実現することを目的としている。

#### 【0010】

##### 【課題を解決するための手段】

本発明は、複数の画像・音声を入力とし、リアルタイムでMPEG-4ビデオオブジェクト(V0)への符号化・編集・配信を一体的に行う実時間コンテンツ編集システムを提供する。本システムは、離れた複数地点からのデジタルビデオによるライブ映像を用いたコンテンツを提供することが可能である。

#### 【0011】

本システムは、編集・配信・受信の3種類のモジュールにより構成され、配信・受信の各モジュールはシステムの稼動中に自由に参加・離脱が可能である。また、配信モジュールは使用するマシンの性能により使用するVOPの種類および頻度を変更することにより、圧縮効率の良い符号化法を自動で選択することが可能である。

#### 【0012】

本発明の実時間コンテンツ編集方法及びシステム並びにプログラムは、インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する。本発明は、入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の1つを選択して符号化可能な符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少

なくとも一つの編集モジュールとを備えている。配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンVOPの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択することを特徴としている。

#### 【0013】

##### 【発明の実施の形態】

###### (実時間コンテンツ編集システム構成)

実時間コンテンツ編集システムは、インターネット上に散在する、ライブ映像を含む多数の画像・音声を素材にした編集処理を行い、複数のユーザに対して配信するシステムである。本システムは、既存のさまざまな環境に柔軟に対応するため、各種回線・マシンによらず動作することを目指している。そのため、画像の符号化として、高い圧縮率を持ちさまざまな帯域に対応したMPEG-4規格を用いる。さらに、複数の画像・音声の編集のために、シーン記述言語を使用する。

#### 【0014】

図1に、構築するシステムの概念図を示す。本システムは、入力装置として複数のビデオカメラと、それらによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールから構成される。また、例示のシステムは、1つの配信サーバと、1つの編集サーバと、2つのクライアントがインターネットを介して接続されるものとして示している。編集モジュールは、編集サーバに配置され、かつ、配信モジュールは、配信サーバに配置するだけでなく、編集サーバにも配置されるものとして例示している。受信モジュールは、クライアントのそれぞれに配置され、かつ、一方のクライアントには、受信モジュールに加えて、配信モジュールも配置されるものとして例示している。図中、DV (Digital Video) ストリーム、MPEG-4 V0、マルチキャスト要求、入出力要求、シーン記述は、それぞれ線の種類を異にする矢印によって示している。

#### 【0015】

編集モジュールはオペレータからの指示に従い配信サーバに対する画像・音声のマルチキャスト要求を実際にを行い、クライアントへ送信するシーン記述言語の

生成およびマルチキャストを担当する。配信モジュールは、入力装置からDV (Digital Video) 形式で取り込まれた画像・音声のMPEG-4符号化を行い、複数のクライアントに対しインターネットを介してパーソナライズメディアストリーム配信によるマルチキャストを行う。受信モジュールを持つクライアントでは、複数のMPEG-4画像音声データおよびシーン記述に基づいて動画像を表示する。ビデオカメラと、配信モジュールを含む装置は、IEEE1394ポートを介して接続される。IEEE1394ポートは、多くのDV機器に備えられており、高速な上に一定周期でデータを送受信するアイソクロナス転送をサポートしているため、映像音声データの送受信に適している。

#### 【0016】

配信モジュール、受信モジュール、編集モジュールとしては任意の形態が考えられる。例えば、汎用のサーバマシン上のプログラムとして配置することもできるし、家庭用PC上で動作させても構わない。また、PCに対する組み込みシステムとして、専用のPCカードの形態を取ることも考えられる。

#### 【0017】

(配信モジュールにおけるDV-MPEG変換器)

##### 1) 構成

実時間処理の可能なコンテンツ編集システムでは、送信者側からデジタルビデオカメラを使用して手持ちのPCを配信サーバとして画像を配信することも、職場の会議室などから高スペックのサーバマシンを使用して配信することも可能である。前者の場合、後者に比べて低スペックなPCを使用することが多く、複雑な符号化処理を行うのは困難である。後者の場合、より複雑な符号化処理が可能であるが、サーバマシンのスペックにより利用可能な符号化の種類が異なる。また、送信者が使用する回線の帯域による制限も存在する。送信側でADSLのような非対称な回線を使用していた場合、上り回線の帯域が数百キロbpsと、下り回線に比べて大きく制限される。以上のような理由から、配信側の実情に合った符号化を行いう必要がある。

#### 【0018】

多くの場合、動画質の維持の観点から一定以上のフレームレートを保つことが

求められる。そのため本システムではR(*frames/s*)（R：所定数）以上のフレームレートを維持するという拘束条件が課せられるものとする。

#### 【0019】

I-VOPのみを使用して符号化を行う場合、フレーム間圧縮に関わる部分である動き補償が不要になるため、直前にエンコードした画像が不要になる。したがって、符号化後の画像に対する逆量子化や逆DCT（離散コサイン変換）を省略することができるため、大幅に符号化時間を短縮することができる（非特許文献6参照）が、符号化後の符号量が大きくなることや任意形状の符号化が行えないことが問題となる。なお、I-VOPとは、入力画像をブロックに分け、ブロック毎にDCTを施してから量子化し、この量子化したDCT係数を可変長符号化（Intra符号化）したVOPを表している。

#### 【0020】

符号量が大きくなる問題については、MPEG-4規格では量子化スケールを設定できるため符号量の調節が可能だが、量子化スケールを大きく設定するに従って画質が悪化するため、変換器を搭載するホストのCPUパワーに余力がある場合にはフレーム間圧縮も行うことが望ましい。そのため、フレーム間圧縮の可能なP-VOP（前方予測符号化されたVOP）やB-VOP（双方向予測符号化されたVOP）を含めた符号化の可能な変換器を構築する。

#### 【0021】

図2に、DVデータをMPEG符号化データに変換する変換器の構成を示す。この変換器構成自体、及びその動作は周知である。本発明は、符号化処理を行うに際して、図2中のすべての符号化処理を、基本処理と付加的処理に分類する。うち基本処理とは最低限必要な符号化処理、付加的処理は処理時間に余裕がある時のみ行う処理である。図2中、太枠で示された各処理、「可変長復号」「逆量子化」「逆DCT」「再構成」「DCT」「量子化」「可変長符号化」は基本処理に該当する。これらは、全VOPをI-VOPとした時に必要となる最小限度の処理である。これに対して、「AC/DC予測」「逆量子化」「逆DCT」「動き補償」「動き情報符号化」は付加的処理と定義される。これらのうち「AC/DC予測」はI-VOPおよびP-VOPに対して符号量を節約するために用いられ、他の各処理はP-VOPおよびB-VOPに

によるフレーム間圧縮のために用いられる。

【0022】

2) 配信モジュールにおける符号化処理

例示のシステムでは、P-VOPやB-VOPを使用するか否かと、それらの使用頻度を決定するため、各種処理の符号化実行時間を計測して1VOPあたりの平均符号化時間を予測し、定められたアルゴリズムに従って判定を行う。図3は、このような配信モジュールにおける符号化処理を説明する図である。

【0023】

ステップS1：

上述したように、符号化処理を、基本処理と付加的処理に分類する。

【0024】

ステップS2：

配信モジュール起動時にIBPBIBPBIBPB…の順に入力画像の符号化を行って各種VOPについて複数個分の処理時間計測を行い、各種処理に要する平均処理時間を測定する。ここで、I、B、Pはそれぞれ、I-VOP、B-VOP、P-VOPをそれぞれ表している。

【0025】

この処理時間計測では、DVデコーダにおける「可変長復号」「逆量子化」「逆DCT」「ブロック再構成」、MPEG-4エンコーダにおける「DCT」「量子化」「可変長符号化」「AC/DC予測」「逆量子化」「逆DCT」「動き補償」「動き情報符号化」の1VOPあたりの処理時間を計測して平均値を求める。これらのうち、「動き補償」はP-VOPとB-VOPで異なる処理を行っており処理時間が大きく異なるため、それぞれ別個に平均値を求める。このようにして、各種処理の符号化実行時間を計測する。

【0026】

各VOPの符号化に必要な時間について、さらに説明する。全VOPの符号化をI-VOPのみを使用して実行する場合、最短の時間で符号化することが可能であるが、P-VOPやB-VOPを使用した符号化を行う場合、前後のVOPと動き補償を行うための参照画像を用意する必要があるため、逆量子化や逆DCTが必要になる。したがって

、同一VOP内でP-VOPやB-VOPを使用するか否かにより、I-VOPの符号化に必要な時間が異なる。評価に必要な各パラメータを以下のように定義する。

$T_{min}$ :最小限の符号化に必要とされる時間

$T_{ACDC}(n)$ :n番目のVOPのAC/DC予測に要する時間

$T_{IQ}$ :逆量子化に必要な時間

$T_{IDCT}$ :逆DCTに必要な時間

$T_c$ :動き情報符号化に要する時間

$T_M(n)$ : n番目のVOPの動き補償に要する時間

#### 【0027】

上記のうち、 $T_M$ と $T_{ACDC}$ を除く各数値は使用するマシンのスペックや対象画像の大きさに対して固定値を取る。これに対し、途中で動き補償の処理を中断しても中断結果を用いた符号化が可能なため、 $T_M$ は可変値に設定することが可能である。また、AC/DC予測はフレームごとに行う／行わないの決定が可能なため、やはり可変値を取る。具体的には、 $T_M$ はP-VOPでは $0 \leq T_M(n) \leq T_{M-Pmax}$ 、B-VOPでは $0 \leq T_M(n) \leq T_{M-Bmax}$ のいずれかの値、 $T_{ACDC}$ は、0もしくは $T_{ACDCmax}$ のいずれかの値を取る。ここで、 $T_{M-Pmax}$ 、 $T_{M-Bmax}$ 、 $T_{ACDCmax}$ はそれぞれ、P-VOPおよびB-VOPにおいてすべてのマクロブロックに対する動き補償を行う際に要する時間、およびAC/DC予測を実行した場合に1VOPあたりのAC/DC予測に要する時間と定義される。上記の各パラメータは、以下の処理時間を計測し、その平均値を算出することにより得られる。

#### 【0028】

$T_{min}$ : DVデコーダにおける「可変長復号」「逆量子化」「逆DCT」「ブロック再構成」、MPEG-4エンコーダにおける「DCT」「量子化」「可変長符号化」の処理時間の合計

$T_{ACDCmax}$ : 「AC/DC予測」の処理時間

$T_{IQ}$ : MPEG-4エンコーダにおける「逆量子化」の処理時間

$T_{IDCT}$ : MPEG-4エンコーダにおける「逆DCT」の処理時間

$T_c$ : 「動き情報符号化」の処理時間

$T_{M-Pmax}$ : P-VOPにおける「動き補償」の処理時間

$T_{M-Bmax}$ : B-VOPにおける「動き補償」の処理時間

なお、上記のうち「動き補償」は、完全に処理を実行した際の処理時間を計測するものとしている。

### 【0029】

各VOPの符号化時間は、以下のようになる。

I-VOPの符号化

- 他のVOPでP-VOPやB-VOPを使用しない場合

$$T(n) = T_{min} + T_{ACDC}(n) \quad (1)$$

- 他のVOPでP-VOPやB-VOPを使用する時

$$T(n) = T_{min} + T_{ACDC}(n) + T_{IQ} + T_{IDCT} \quad (2)$$

P-VOP、B-VOPの符号化

$$T(n) = T_{min} + T_{ACDC}(n) + T_{IQ} + T_{IDCT} + T_c + T_M(n) \quad (3)$$

N個のVOP中に出現するI-VOP、P-VOP、B-VOPの数をそれぞれNIVOP、NPVOP、NB  
VOPと置くと、N個のVOPにおける符号化時間の合計は

$$\begin{aligned} T_{total} &= \sum_{I-VOP} T(n) + \sum_{P-VOP} T(n) + \sum_{B-VOP} T(n) \\ &= N_{IVOP} (T_{min} + T_{IQ} + T_{IDCT}) + \sum_{I-VOP} T_{ACDC}(n) \\ &\quad + N_{PVOP} (T_{min} + T_{IQ} + T_{IDCT} + T_c) + \sum_{P-VOP} (T_{ACDC}(n) + T_M(n)) \\ &\quad + N_{BVOP} (T_{min} + T_{IQ} + T_{IDCT} + T_c) + \sum_{B-VOP} (T_{ACDC}(n) + T_M(n)) \\ &= N \times (T_{min} + T_{IQ} + T_{IDCT}) + (N_{PVOP} + N_{BVOP}) T_c \\ &\quad + \sum_{all-VOP} T_{ACDC}(n) + \sum_{P-VOP} T_M(n) + \sum_{B-VOP} T_M(n) \end{aligned} \quad (4)$$

その際の1フレームの平均符号化時間は

$$\begin{aligned} T_{ave} &= T_{total} / N = T_{min} + T_{IQ} + T_{IDCT} + T_c (N_{PVOP} + N_{BVOP}) / N \\ &\quad + (\sum_{all-VOP} T_{ACDC}(n) + \sum_{P-VOP} T_M(n) + \sum_{B-VOP} T_M(n)) / N \end{aligned} \quad (5)$$

$T_M$ 、 $T_{ACDC}$ の各値が可変であることから、(5)式のうち、 $\Sigma$ で示された各項目を自由に定めることができる。 $T_{ACDCmax}$ は $T_{M-Pmax}$ の数分の一程度であり、 $T_M$ - $Pmax$ は $T_{M-Bmax}$ の数分の一程度であることから、AC/DC予測、P-VOPの動き補償、B-VOPの動き補償の順に優先的に符号化時間を配分する。実際にはAC/DC予測の符号化全体に占める時間割合は小さな値のため、通常はAC/DC予測を実行するものとして説明する。

## 【0030】

ステップS3：

ステップ2で測定された平均処理時間の値を用いて、VOP判定アルゴリズムに従って使用するVOPの種類および頻度を決定する。

## 【0031】

目標フレームレートをRと置くと、1フレームの目標平均符号化時間は $1/R$ となる。(5)式より、N個のVOP中において $N_{PVOP}$ 個のP-VOPを使用して完全に動き補償を実行する際の平均符号化時間は、

$$T_{ave} = T_{min} + T_{IQ} + T_{IDCT} + (T_c + T_{M-Pmax}) N_{PVOP}/N + T_{ACDCmax} \quad (6)$$

となる。したがって、N個のVOP中において $N_{PVOP}$ 個のP-VOPを使用するための最低条件は

$$1/R \geq T_{min} + T_{IQ} + T_{IDCT} + (T_c + T_{M-Pmax}) N_{PVOP}/N + T_{ACDCmax} \quad (7)$$

となる。また、N個のVOP中において $N_{PVOP}$ 個のP-VOPおよび $N_{BVOP}$ 個のB-VOPを使用して、P-VOPおよびB-VOP中のすべてのマクロブロックに対して完全に動き補償を行った場合、平均符号化時間は

$$\begin{aligned} T_{ave} = & T_{min} + T_{IQ} + T_{IDCT} + (T_c + T_{M-Pmax}) N_{PVOP}/N \\ & + (T_c + T_{M-Bmax}) N_{BVOP}/N + T_{ACDCmax} \end{aligned} \quad (8)$$

となる。したがって、N個のVOP中において $N_{PVOP}$ 個のP-VOPおよび $N_{BVOP}$ 個のB-VOPを使用するための条件は

$$\begin{aligned} 1/R \geq & T_{min} + T_{IQ} + T_{IDCT} + (T_c + T_{M-Pmax}) N_{PVOP}/N \\ & + (T_c + T_{M-Bmax}) N_{BVOP}/N + T_{ACDCmax} \end{aligned} \quad (9)$$

となる。

## 【0032】

実際には各VOPの比率は自由に変更できるわけではなく、2個のI-VOP間に挟まれるP-VOPの数や、2個のI-VOPもしくはP-VOPの間に挟まれるB-VOPの数を変更することにより比率が決定される。連続する2個のI-VOPの間に挟まれるP-VOPの数に1を加えたものを $N_p$ 、連続する2個のI-VOPもしくはP-VOPの間に挟まれるB-VOPの数に1を加えたものを $N_b$ と定義すると、N個の連続したVOPに含まれる各VOPの数は以下のようになる。

$$N_{IVOP} = N / N_p \cdot N_b \quad (10)$$

$$N_{PVOP} = (N_p - 1) \cdot N / N_p \cdot N_b \quad (11)$$

$$N_{BVOP} = (N_b - 1) \cdot N / N_b \quad (12)$$

このようにすると、(7)式および(9)式は以下のようになる。

$$\begin{aligned} 1/R \geq & T_{min} + T_{IQ} + T_{IDCT} + (T_c + T_{M-Pmax}) \cdot (N_p - 1) / N_p \cdot N_b \\ & + T_{ACDCmax} \end{aligned} \quad (13)$$

$$\begin{aligned} 1/R \geq & T_{min} + T_{IQ} + T_{IDCT} + (T_c + T_{M-Pmax}) \cdot (N_p - 1) / N_p \cdot N_b \\ & + (T_c + T_{M-Bmax}) \cdot (N_b - 1) / N_b + T_{ACDCmax} \end{aligned} \quad (14)$$

適正なNBVOP、NPVOPを決定するためのアルゴリズムを図4に示す。アルゴリズムの入力TおよびRは、それぞれ各符号化処理の処理時間および目標フレームレートである。うちTは、全ての符号化処理時間を含み、 $T = [T_{min}, T_{ACDC}, T_{IQ}, T_{IDC}, T_c, T_{M-Pmax}, T_{M-Bmax}]$ で示される。

#### 【0033】

始めに、 $N_{BVOP} = 1$ 、 $N_{PVOP} = 2$ で(13)式の判定を行う。これで全VOPの動き補償が可能なようなら $N_{PVOP}$ を二倍ずつ増やして判定する。これを、あらかじめ設定されている最大値である $N_{PVOPmax}$ に達するまで判定を続ける。この途中で(13)式を満たさない条件が得られた場合、 $N_{PVOP}$ を直前の値に戻して、I-VOPとP-VOPを使用した符号化を行う。

#### 【0034】

$N_{PVOP} = N_{PVOPmax}$ に達してもなお全VOPの動き補償が可能なようなら $N_{BVOP} = 2$ とし、以後二倍ずつ増やして(14)式の判定を行い、 $N_{BVOPmax}$ に達するまで判定を続ける。この途中で(14)式を満たさない条件が得られた場合、 $N_{BVOP}$ を直前の値に戻して、全種類のVOPを使用した符号化を行う。その際、 $N_{BVOP}$ を増やしてゆくに従ってI-VOPが減るので、I-VOPの比率を一定に保つため、 $N_{BVOP}$ を二倍にするたびに $N_{PVOP}$ を二分の一にする。

#### 【0035】

ステップS4：

ステップS3で決定されたVOPの種類及び頻度に従って符号化を続けてゆく。

#### 【0036】

【発明の効果】

本発明は、複数の画像・音声を扱った編集・配信作業が可能なシステムを、編集・配信・受信の3種類のモジュールにより構成し、その配信モジュールを使用するマシンの性能に応じて使用するVOPの種類を変更することにより、可能な限り圧縮効率の良い符号化法を自動で選択することが可能となる。

【図面の簡単な説明】

【図1】

構築するシステムの概念図を示す。

【図2】

変換器の構成を示す図である。

【図3】

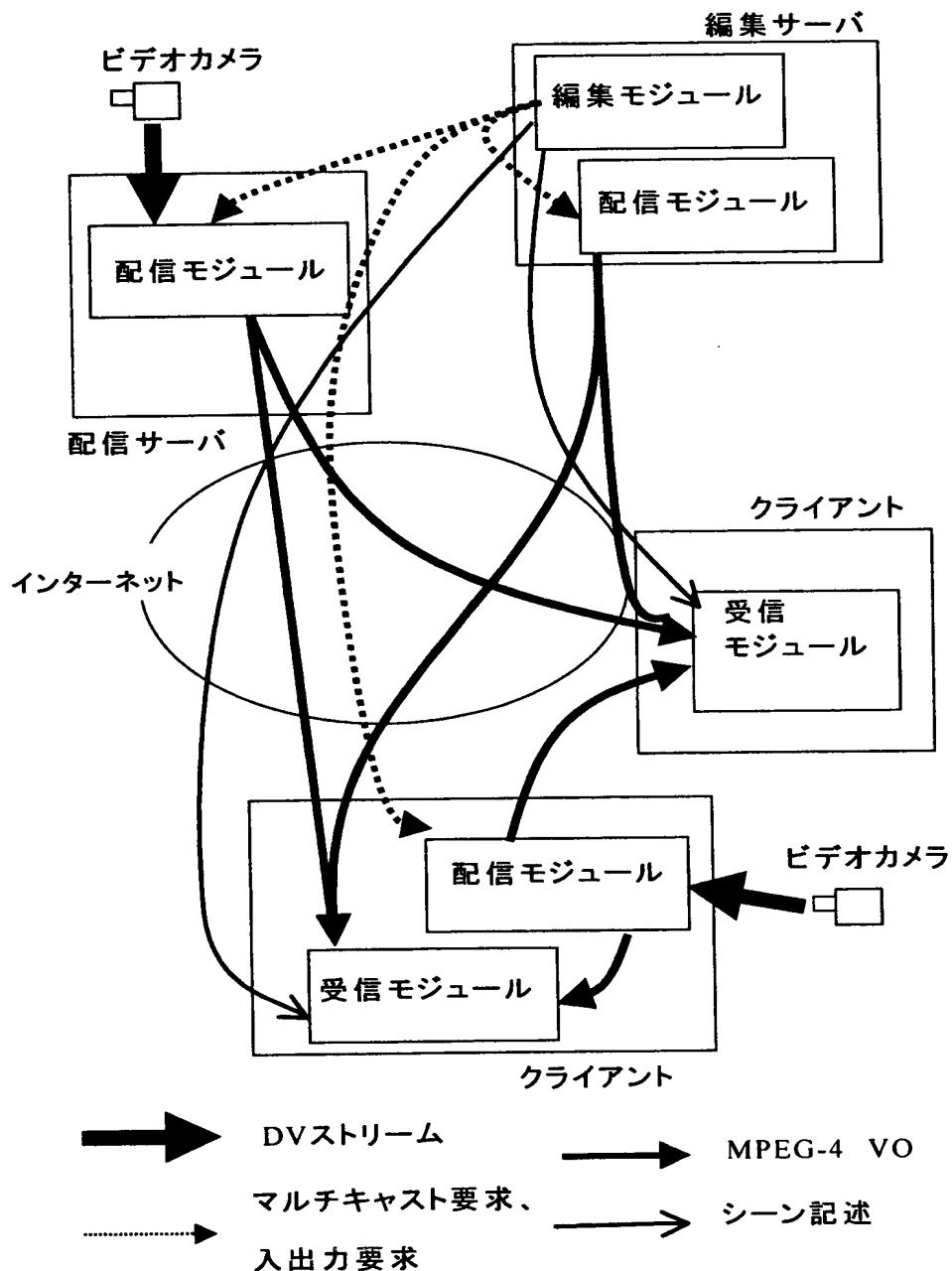
配信モジュールにおける符号化処理を説明する図である。

【図4】

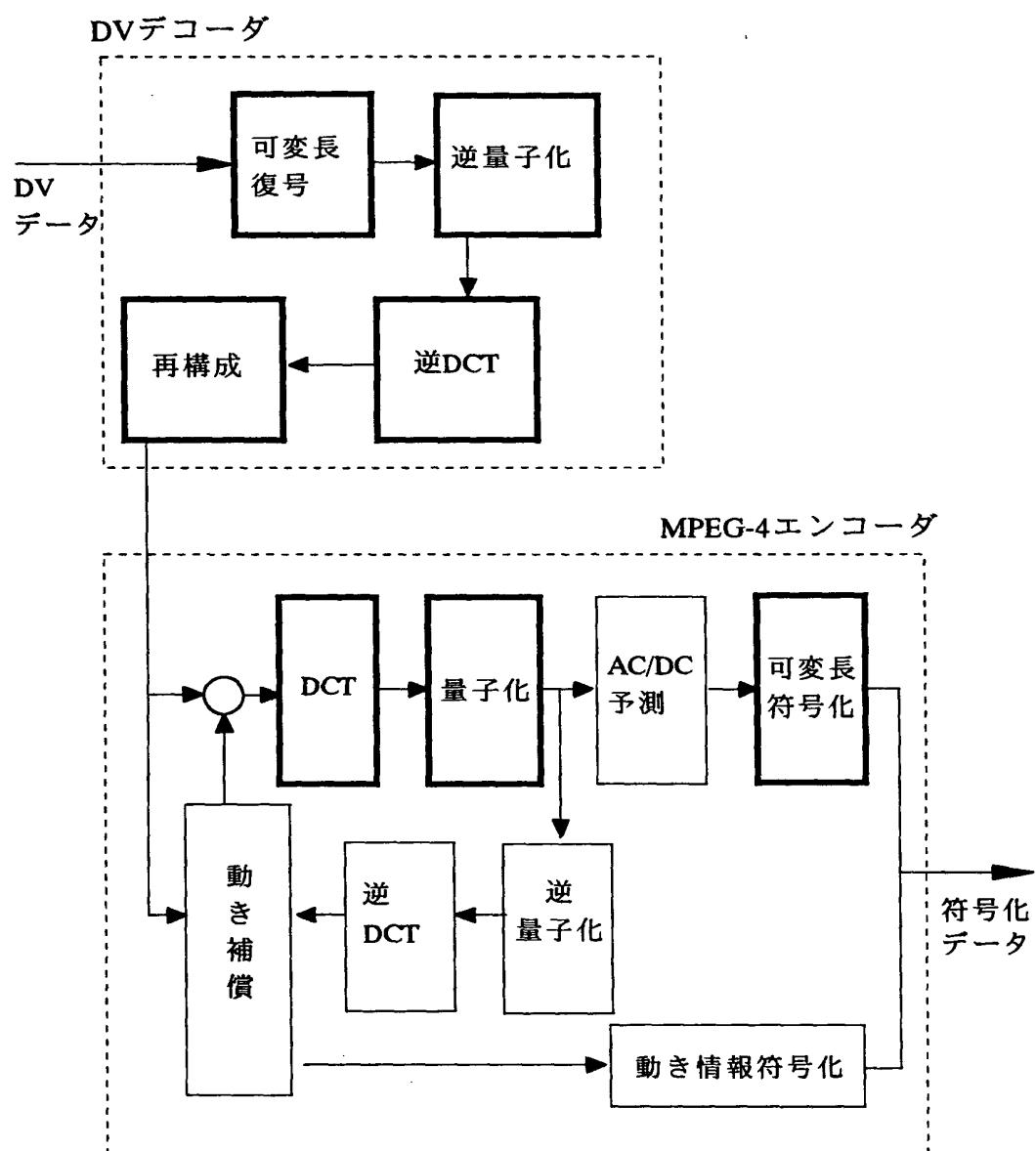
適正なNBVOP、NPVOPを決定するためのアルゴリズムを示す図である。

【書類名】図面

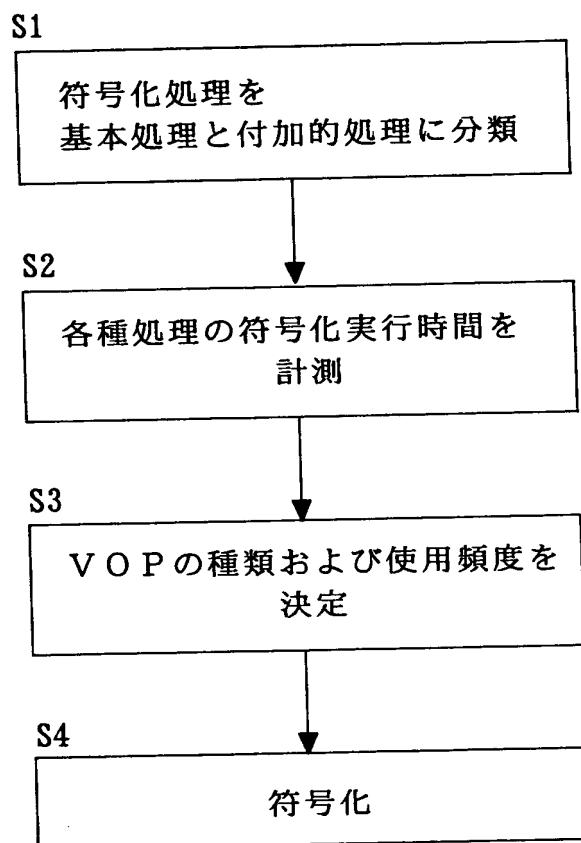
【図1】

実時間コンテンツ編集システムの構成例

【図2】

DV-MPEG変換器の構成

【図3】

配信モジュールにおける符号化処理

【図4】

符号化パラメータ決定のためのアルゴリズム

```

Decide_param(T, R){
    Nb=1;
    for(Np=2;Np<=Npmax;Np=Np*2)
        if(1/R < Tmin+TIQ+TIDCT
           +(Tc+TM*Pmax)(Np-1)/NpNb+TACDCmax){
            Np=Np/2;
            return(Np,Nb);
        }
    Np=Npmax;
    for(Nb=2;Nb<=Nbmax;Nb=Nb*2){
        Np=Np/2;
        if(1/R < Tmin+TIQ+TIDCT
           +(Tc+TM*Pmax)(Np-1)/NpNb
           +(Tc+TM*Bmax)(Nb-1)/Nb+TACDCmax){
            Nb=Nb/2;
            return(Np,Nb);
        }
    }
    Nb=Nbmax;
    Np=Npmax/Nbmax;
    return(Np,Nb);
}

```

【書類名】 要約書

【要約】

【課題】 CPUパワーに応じた符号化処理を行うことで、リアルタイムなMPEG-4 V0への符号化を実現する。

【解決手段】 本発明は、インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する。本発明は、入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の1つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えている。配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンVOPの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する。

【選択図】 図1

出願人履歴情報

識別番号 [301022471]

1. 変更年月日 2001年 4月 2日

[変更理由] 新規登録

住 所 東京都小金井市貫井北町4-2-1  
氏 名 独立行政法人通信総合研究所